

# RNA-Bibliothekserstellung mit Agilent SureSelect XT HS: Ein optimierter und verbesserter Arbeitsablauf für die Direct-to-Capture-Erstellung von RNA-Sequenzierbibliotheken aus frischen oder FFPE-Proben

## Autoren

Carsten Carstens, Katherine Felts und Sarah Johns  
Agilent Technologies, Inc.

## Zusammenfassung

In dieser Application Note stellen wir einen komprimierten und verbesserten Arbeitsablauf für die gezielte Erstellung von RNA-Sequenzierbibliotheken vor. Das Agilent SureSelect XT RNA Direct Protokoll wurde in vier wesentlichen Punkten verbessert:

1. Ersatz der Übernacht-Hybridisierung zur Target-Anreicherung durch eine 90-minütige Schnellhybridisierung
2. Wegfall der Behandlung mit Uracil-Deglycosylase (UDG) (RNA-Strangspezifität wird durch ein neues Enzym gewährleistet)
3. Austausch der älteren SureSelect XT Adapter durch SureSelect XT HS Adapter, welche die gleichzeitige Verarbeitung von DNA und RNA aus der gleichen Probe ermöglichen
4. Bereitstellung eines eindeutigen molekularen Barcodes (molecular barcode, MBC) für die verbesserte Deduplizierung von PCR- und Fragmentierungsduplikaten

Durch diese Änderungen konnte ein Protokoll mit einer Bearbeitungszeit von 2–3 Tagen auf nur einen Tag reduziert werden. Zusätzlich fanden wir durch die Analyse von Fusionsdaten bei Verwendung verschiedener Proben typen (intakt, frisch gefroren, FFPE) heraus, dass dieser optimierte Arbeitsablauf bei einem Input von nur 10 ng RNA RNA-Sequenzierbibliotheken mit höherer Bibliothekskomplexität und eine verbesserte Sequenzierleistung hervorbringt.

## Einleitung

Die Anwendung der Hochdurchsatz-Sequenzierung auf Transkriptomik (RNA-Sequenzierung, RNA-Seq) hat nicht nur das globale Genexpressions-Profilierung ermöglicht, sondern auch präzise Informationen zu Spleißvarianten, Fusionstranskripten, posttranslationalen Modifikationen und allelspezifischer Expression geliefert. Der Standardansatz zur Erstellung von RNA-Seq-Bibliotheken umfasst die Fragmentierung

von RNA durch Mg<sup>2+</sup>-abhängige Umesterung, die durch zufällige Hexamere initiierte Synthese des ersten cDNA-Strangs sowie die Synthese des zweiten cDNA-Strangs mit dUTP-Strangmarkierung. Zur Vermeidung von Verknüpfungen erhält die cDNA stumpfe Enden, und den 3'-Enden wird durch nicht-templategebundene DNA-Polymerisierung ein Adenin hinzugefügt. Anschließend folgt die Ligation plattformspezifischer Sequenzieradapter. Alle Schritte nach der Adapter-Ligation sind mit dem Arbeitsablauf einer DNA-Sequenzierung identisch.

Obwohl DNA- und RNA-basierte Sequenzierbibliotheken nach der Adapter-Ligation identisch erscheinen, sind beim Vergleich von RNA- und DNA-Sequenzierung bestimmte Unterschiede zu beachten. Der erste Unterschied besteht in der Notwendigkeit, die Direktionalität der Sequenzierfragmente aufrechtzuerhalten, um feststellen zu können, welchem genomischen DNA-Strang die ursprüngliche RNA entspricht. Dies wird üblicherweise erreicht, indem der Reaktion zur Synthese des zweiten Strangs Uracil beigefügt, die Bibliothekserstellung fortgesetzt und der zweite Strang in einem späteren Schritt durch Zugabe von UDG entfernt wird. Der zweite Unterschied ist der große dynamische Bereich der bei der RNA-Sequenzierung gemessenen Transkripte. Er kann sich aufgrund der bedeutenden Unterschiede in der relativen Häufigkeit exprimierter Transkripte und der extrem hohen Abundanz ribosomaler RNA (rRNA) verglichen mit codierender RNA über mehr als fünf Größenordnungen erstrecken. Für die RNA-Seq ist daher eine Reduzierung der Komplexität erforderlich, um zu vermeiden, dass Reads für Transkripte ohne Aussagekraft (z. B. rRNA) verschwendet werden. Bei der häufigsten Methode wird rRNA entweder durch gezielte Depletion (Ribodepletion) oder die spezifische Bindung von polyadenylierter RNA (mRNA) entfernt. In einer typischen Probe macht selbst nach rRNA-Entfernung das 1 % der am häufigsten exprimierten Gene noch immer ≈ 50 % aller Transkripte aus. Wenn Forschungsziele also die Untersuchung von schwach oder mittelstark exprimierten Genen erfordern, kann deren Detektion durch eine weitere Komplexitätsreduktion erleichtert werden.

Als Alternative zu gezielten Depletionsansätzen kann eine Reduktion der Komplexität auch durch die gezielte Anreicherung mithilfe biotinylierter Target-Anreicherungs sonden (sog. „Baits“) erreicht werden. Dieses Vorgehen ist in der genomischen Sequenzierung weit verbreitet. Im Rahmen der RNA-Sequenzierung wird die gezielte Anreicherung durch Bait-Selektion hauptsächlich für FFPE-Proben eingesetzt, da bei diesen die rRNA-Depletion bekanntermaßen inkonsistent ist und eine Poly(A)-Anreicherung aufgrund der Fragmentierung des Ausgangsmaterials nicht infrage kommt<sup>1</sup>. Die Target-Anreicherung ist auch dann von Vorteil, wenn nur eine relativ kleine Teilmenge des Transkriptoms untersucht werden muss. Ein prominentes Beispiel ist die Detektion von Fusionstranskripten, die auf eine zugrundeliegende Genfusion hindeuten. Allerdings profitiert von der gezielten Anreicherung praktisch jedes Szenario, in dem nur eine Untermenge der Transkripte informativ ist, z. B. transkriptionelles Fingerprinting oder Detektion seltener posttranslationaler Modifikationen<sup>2</sup>.

Agilent bietet das SureSelect XT RNA Direct Kit zur Bibliothekserstellung (Best.-Nr. G7564A, G7564B) an, welches die Konstruktion gezielter RNA-Sequenzierbibliotheken ermöglicht. Wir haben den erfolgreichen Einsatz dieses Kits bei der Arbeit mit FFPE-Proben demonstriert<sup>3</sup>. Hier beschreiben wir einen optimierten und verbesserten Arbeitsablauf, der durch Kombination von Komponenten des SureSelect XT RNA Direct Kits zur Bibliothekserstellung mit dem SureSelect XT HS Kit zur Target-Anreicherung (Best.-Nr. G9706A) in Verbindung mit einem modifizierten Protokoll zustande kommt. Durch Eliminierung der Lyophilisierung der Input-RNA, Eliminierung eines Aufreinigungsschritts mit SPRI-Beads und Ersatz der 24-Stunden-Bait-Hybridisierung durch einen 1,5-stündigen Schnellhybridisierungsschritt wurde der Arbeitsablauf vereinfacht. Zusätzlich wurde die Behandlung mit UDG eliminiert. Die Strangspezifität wird nun stattdessen mithilfe eines PCR-Enzyms gewährleistet, das im Rahmen der Precapture-PCR uracilhaltige Template-DNA erkennt und diese nicht

amplifiziert. Durch die Verwendung von SureSelect XT HS Sequenzieradaptern wird außerdem ein MBC eingebracht, der die Identifizierung von Fragmentierungsduplikaten ermöglicht und die gleichzeitige („Split Tube“, also parallele) Verarbeitung von DNA und RNA aus der gleichen Probe vereinfacht.

## Experimentelles

### RNA-Quellen

Die universelle Human-Referenz-RNA (UHRR) wurde als frisch gefrorenes Material von Agilent Technologies (Santa Clara, CA, USA, Best.-Nr. 750500-41) bezogen. Ein Satz aus Brusttumor- und angrenzendem Normalgewebe wurde in Form von frisch gefrorenem und FFPE-Material von CureLine Human Biospecimen CRO (Brisbane, CA, USA, kundenspezifische Bestellnummer) bezogen. Seraseq FFPE Tumor Fusion RNA Reference Material v2 wurde von SeraCare (Gaithersburg, MD, USA, Kat.-Nr. 0710-0129) bezogen.

### RNA-Isolierung

Soweit erforderlich, wurde RNA mit dem RNeasy FFPE Kit oder dem RNeasy Mini Kit von Qiagen (Qiagen USA, Germantown, MD, USA, Best.-Nr. 73504 bzw. 74104) gemäß den Anweisungen des Herstellers isoliert. Ein ausführlicheres Protokoll ist im Anhang zu finden.

### Qualitätsbewertung des Einsatzmaterials und der Sequenzierbibliotheken

Nukleinsäureproben wurden auf dem Agilent 2100 Bioanalyzer System (Agilent Technologies, Best.-Nr. G2939B) entweder mit dem Agilent RNA 6000 Pico Kit (Agilent Technologies, Best.-Nr. 5067-1513) zur Beurteilung der RNA-Qualität oder dem Agilent DNA 1000 Kit (Agilent Technologies, Best.-Nr. 5067-1504) zur Bewertung der Qualität der Sequenzierbibliotheken untersucht.

### Andere Materialien

Actinomycin D wurde von Sigma (St. Louis, MO, USA, Best.-Nr. A1410) bezogen und in DMSO-Stammllösung auf 4 µg/µl eingestellt. SPRI-Bead-Aufreinigungen wurden mit AMPure XP Beads (Beckman Coulter, Atlanta, GA, USA,

Best.-Nr. A63880) vorgenommen. Für die Bindung biotinylierter Sonden wurden DynabeadsMyOne Streptavidin T1 Beads (Thermo Fisher Scientific, Waltham, MA, USA, Kat.-Nr. 65601) verwendet.

### SureSelect XT HS RNA Bibliothekserstellung

Die Konstruktion von RNA-Sequenzierbibliotheken wurde mithilfe des SureSelect XT RNA Direct Kits (Agilent Technologies, Best.-Nr. G7564A) und des SureSelect XT HS Systems zur Target-Anreicherung für die Illumina Paired-End-Multiplex-Sequenzierbibliothek (Agilent Technologies, Best.-Nr. G9706A) durchgeführt. Eine ausführliche Beschreibung ist im Anhang zu finden.

### SureSelect XT RNA Direct Bibliothekserstellung

Erstellung, Anreicherung und Sequenzierung der RNA Direct Bibliotheken wurden anhand des SureSelect XT RNA Direct Protokolls zur Bibliothekserstellung (Handbuch) vorgenommen.

### Target-Anreicherung

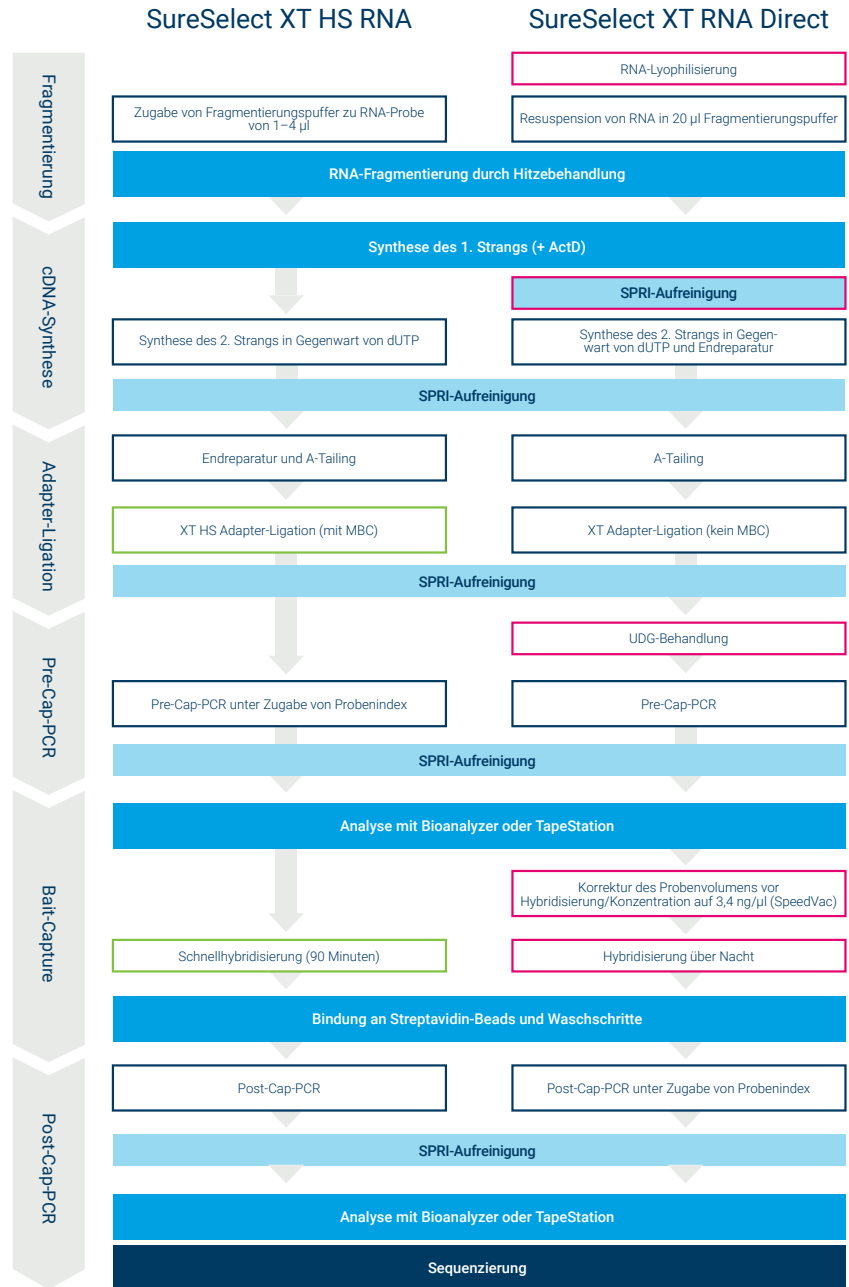
Die Target-Anreicherung für SureSelect XT HS RNA Bibliotheken wurde unter Verwendung des SureSelect Human All Exon V7 Exoms (Agilent Technologies, Best.-Nr. 5191-4029), welches auf das codierende Transkriptom abzielt, durchgeführt. Ein detailliertes Protokoll für das Bait-Capture-Verfahren ist dem Anhang zu entnehmen.

### Sequenzierung und Datenanalyse

Die Sequenzierbibliotheken wurden auf einem Illumina HiSeq 4000 durch Paired-End-Sequenzierung in einem Format von 2 x 150 Reads analysiert. Für die Expressionsanalyse (Daten nicht gezeigt) wurden FASTQ-Dateien mit dem Transkriptom aligniert. Dazu wurde das STAR Paket in der Version 2.6.0a (mit Berücksichtigung des Spleißens) mit Genomversion hg38 als Referenz verwendet. Die Expressionsprofile wurden dann mithilfe des RSEM-Tools aus den Ergebnissen des STAR-Alignments generiert. Die allgemeine Bibliotheksstatistik (Strangspezifität, 5'-3'-Bias, MBC-blinde Duplikationsraten, Schätzungen der Bibliotheksgröße)

wurde unter Anwendung der Picard RNA-Analysepipeline generiert. Duplikate wurden mithilfe von .bam-Dateien markiert, welche auf  $2 \times 10^7$  Reads downgesampelt wurden, um normalisierte Duplikationsraten zu erhalten. Die MBC-korrigierte Duplikationsstatistik sowie die Schätzungen der Bibliotheksgröße wurden anhand der gleichen Pipeline

generiert. Zur Entfernung von Fragmentierungsduplikaten wurde UmiAwareMarkDuplicatesWithMateCigar eingesetzt. Fusionstranskripte wurden mit STAR-Fusion beurteilt und mit dem FusionInspector Tool visualisiert, einem Bestandteil des Trinity Cancer Transcriptome Analysis Tool Kits (CTAT)<sup>4</sup>.



**Abbildung 1.** Vergleich des Agilent SureSelect XT HS RNA Arbeitsablaufs (RNA XT HS) mit dem Agilent SureSelect XT RNA Direct Kit zur Bibliothekserstellung (RNA Direct). Die aus dem RNA Direct Arbeitsablauf eliminierten Schritte sind durch eine pinke Umrandung gekennzeichnet. Schritte, die zu der verbesserten Funktionalität des RNA XT HS Protokolls beitragen, sind durch eine grüne Umrandung gekennzeichnet.

## Ergebnisse und Diskussion

### Überblick über den RNA XT HS und den RNA Direct Arbeitsablauf

Unser Ziel war es, einen vereinfachten und verbesserten Arbeitsablauf für die Erstellung hochwertiger RNA-Sequenzierbibliotheken zu entwickeln. Dies wollten wir durch Kombination von Komponenten des SureSelect XT RNA Direct Kits zur Bibliothekserstellung (RNA Direct) mit dem SureSelect XT HS Kit zur Target-Anreicherung erreichen. Der resultierende Arbeitsablauf wurde als SureSelect XT HS RNA (RNA XT HS) benannt. Ein vergleichender Überblick über diese Arbeitsabläufe ist in Abbildung 1 dargestellt. Der RNA XT HS Arbeitsablauf hat klare Vorteile gegenüber dem RNA Direct Arbeitsablauf, etwa die Eliminierung der UDG-Behandlung und den Ersatz der herkömmlichen Übernacht-Hybridisierung durch eine 90-minütige Schnellhybridisierung, welche die Bearbeitungszeit deutlich reduziert. Wir wollten weiterhin ermitteln, ob der SPRI-Aufreinigungsschritt zwischen der Synthese des ersten und des zweiten Strangs eliminiert sowie die Lyophilisierung des RNA-Inputs im RNA XT HS Arbeitsablauf vermieden werden können (für Details siehe Anhang).

### Leistungsvergleich der Arbeitsabläufe

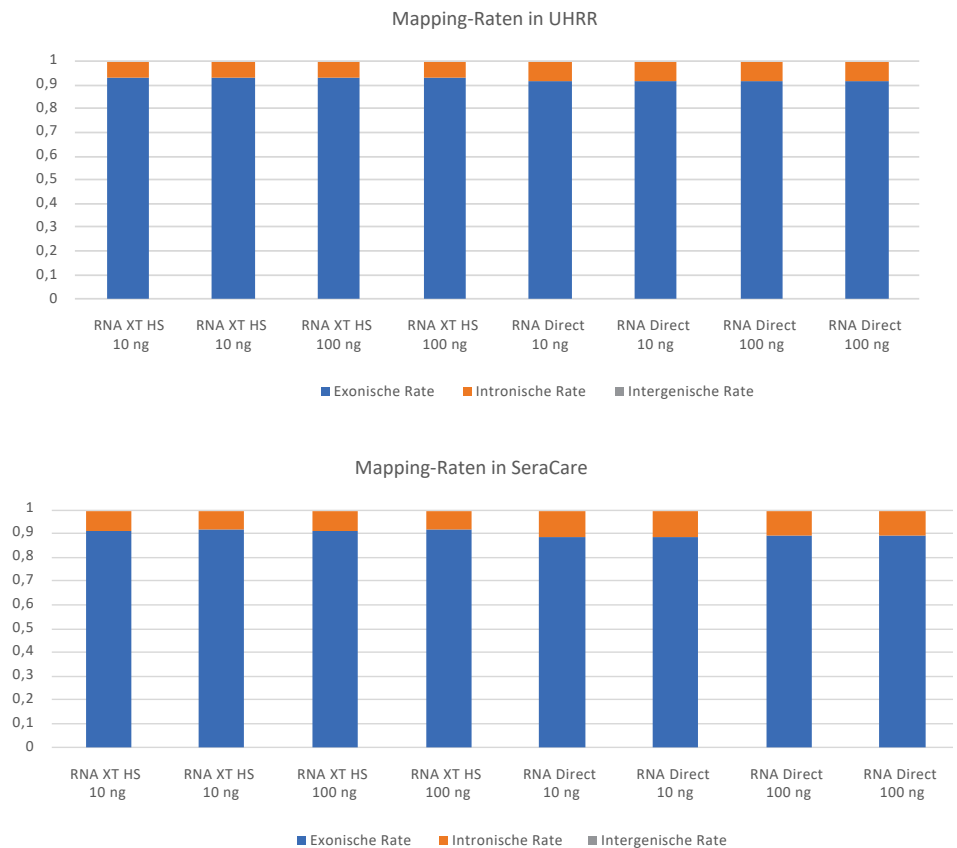
Um die Leistung beider Bibliothekserstellungsprozesse zu vergleichen, generierten wir zunächst Bibliotheken mit verschiedenen Mengen von universeller Human-Referenz-RNA (UHRR) und SeraCare FFPE Tumor Fusion RNA Reference Material v2 (SeraCare) als Einsatzmaterial, repräsentativ für eine intakte Probe bzw. eine idealisierte FFPE-Probe. Die RNA XT HS und RNA Direct Bibliotheken wurden mit dem SureSelect Human All Exon V7 angereichert. Schließlich wurden diese angereicherten Bibliotheken mit einem Illumina-Sequencer sequenziert und die Daten unter Zuhilfenahme individueller Datenanalysepipelines analysiert (für Details siehe Abschnitt „Experimentelles“).

Tabelle 1 zeigt eine Zusammenfassung der globalen Sequenzierstatistikergebnisse aus der Analyse der RNA XT und RNA Direct Sequenzierdaten. Wir stellten

fest, dass die sequenzierten RNA XT HS Bibliotheken hinsichtlich zahlreicher Messgrößen nicht von den RNA Direct Bibliotheken zu unterscheiden sind. Besonders hervorzuheben sind hier die hohe Strangspezifität (> 98 %) und die geringe rRNA-Kontamination (~ 0,1 %). Die hohe Strangspezifität der RNA XT HS Bibliotheken demonstriert die Effektivität des Ansatzes bei RNA XT HS, die UDG-Behandlung durch die Verwendung eines PCR-Enzyms zu ersetzen, das uracilhaltige Templates nicht amplifiziert. Die Strangspezifität für intakte Proben ist tendenziell etwas höher als die für FFPE-Proben, was aufgrund der schlechteren Qualität des FFPE-Einsatzmaterials zu erwarten ist. Trotzdem sind die mit FFPE-Proben als Input beobachteten Strangspezifitäten für beide Protokolle noch immer sehr hoch (> 98 % für FFPE-Material). Der durchweg geringe Prozentsatz an rRNA-Kontamination

für beide Methoden der Bibliothekserstellung ist bereits im Vorfeld in Verfahren zur Target-Anreicherung demonstriert worden<sup>3</sup>.

Beim Vergleich der exonischen Mapping-Raten von RNA XT HS und RNA Direct Bibliotheken ergeben sich erneut vergleichbare Werte (Abbildung 2). Ein kritischer Schritt bei der Verkürzung des neuen RNA XT HS Arbeitsablaufs ist der Austausch der traditionellen Übernacht-Hybridisierung gegen einen Schnellhybridisierungsschritt. Wie Abbildung 2 zeigt, beeinflusst das beschleunigte Capture-Verfahren nicht die Mapping-Raten. Wir konnten durchgängig exonische Raten von 90 % beobachten, während unabhängig vom Einsatzmaterial nur sehr wenige Reads intergenischen Regionen zugeordnet wurden. 92,9 – 94,1 % der gesamten beobachteten Expression entfielen auf von der V7 Capture-Bibliothek angezielte Transkripte, unabhängig

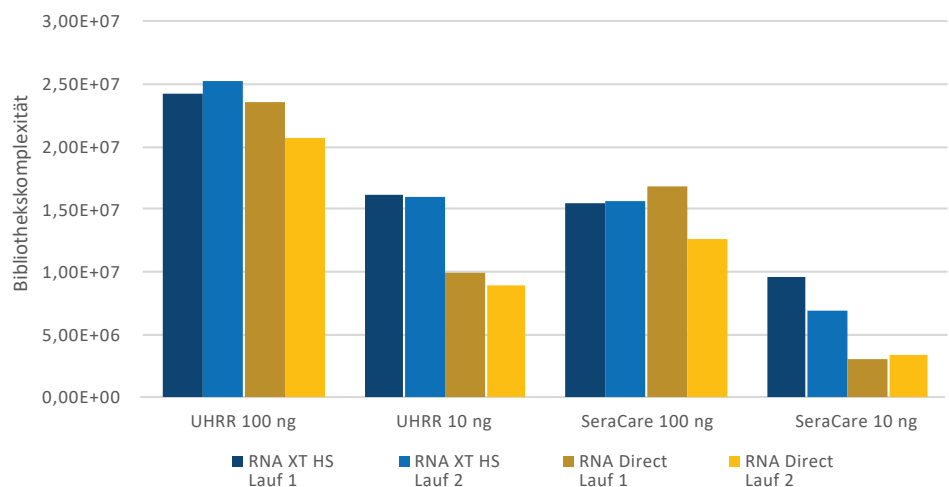


**Abbildung 2.** Vergleich der Mapping-Raten zwischen RNA XT HS und RNA Direct Sequenzierbibliotheken. RNA-Seq-Bibliotheken wurden unter Anwendung des RNA XT HS oder des RNA Direct Arbeitsablaufs aus UHRR- (A) und SeraCare-Proben (B) erstellt. Dargestellt sind die Mapping-Raten der mithilfe des Agilent SureSelect Human All Exon V7 Exoms angereicherten Sequenzierbibliotheken für exonische, intronische und intergenische Sequenzen. Die intronische Rate spiegelt vermutlich die Sequenzierung unprozessierter mRNA wider.

von Einsatzmenge und Probenart. Die verbleibenden ~ 7 % sind größtenteils auf kontaminierende, stark exprimierte Transkripte wie mitochondriale Gene oder auf Annotationsfehler zurückzuführen.

Ein kritischer Indikator für die Effizienz des gesamten Prozesses der Bibliothekserstellung ist die geschätzte anfängliche Bibliothekskomplexität. Größere und damit komplexere Bibliotheken sind vorzuziehen, da die Sequenzierung dieser Bibliotheken genauere Einblicke und eine zuverlässigere Quantifizierung ermöglicht. Die Bibliothekskomplexität wird aus der Sequenziertiefe (in Read-Paaren) und der Anzahl beobachteter einzigartiger Varianten berechnet, die aus der Duplikationsrate abgeleitet wird. Wir haben zur Berechnung der geschätzten Bibliothekskomplexitäten die Picard-Pipeline angewendet, welche alle Duplikate als PCR-Duplikate klassifiziert. Bei diesem Ansatz wird die wahre Bibliothekskomplexität unterschätzt, worauf im Abschnitt „Ergebnisse“ noch weiter eingegangen wird. Die projizierten Bibliothekskomplexitäten sind in Tabelle 1 aufgelistet und in Abbildung 3 dargestellt. Beim Vergleich des RNA XT HS und des RNA Direct Arbeitsablaufs beobachteten wir bei hohen Einsatzmengen (100 ng) frischer oder FFPE-RNA keine signifikanten Unterschiede zwischen den Bibliotheken. Wurde der Input jedoch für UHRR oder SeraCare auf 10 ng reduziert, hatte dies auffallende Unterschiede in der Bibliothekskomplexität zur Folge. Zunächst sind die aus geringeren Einsatzmengen erstellten Bibliotheken kleiner und weniger komplex als jene mit höherem Input. Dies ist zu erwarten, da die fertige Sequenzierbibliothek umso kleiner ausfallen sollte, je weniger RNA zu ihrer Erstellung eingesetzt wird. Wir fanden auch heraus, dass der optimierte RNA XT HS Arbeitsablauf bei geringer Einsatzmenge Bibliotheken generiert, die etwa 1,5- bis 2-mal effizienter ausfallen als mit dem RNA Direct Arbeitsablauf. Verschiedene Faktoren des RNA XT HS Arbeitsablaufs, die zu dieser Steigerung der Effizienz beitragen, werden aktuell untersucht (Daten nicht gezeigt).

Geschätzte Komplexität der Sequenzierbibliotheken (Picard-Pipeline)



**Abbildung 3.** Unterschiede in der Bibliothekskomplexität zwischen RNA XT HS und RNA Direct Bibliotheken bei geringem Probeneinsatz. Unter Verwendung der angegebenen Quellen und Einsatzmengen erstellte Sequenzierbibliotheken wurden auf einer Illumina HiSeq 4000 Plattform mittels Paired-End-Sequenzierung sequenziert. Die geschätzten Bibliothekskomplexitäten wurden auf Grundlage der beobachteten Duplikationsraten und der jeweiligen Anzahl der mit dem Picard RNA-Analyssetool ermittelten Read-Paare bestimmt. Diese Pipeline unterscheidet nicht zwischen PCR- und Fragmentierungsduplikaten.

**Tabelle 1.** Globale Sequenzierstatistik der SureSelect XT HS RNA und SureSelect XT RNA Direct Bibliotheken: Mit dem jeweils gleichen Einsatzmaterial entweder nach dem RNA XT HS oder dem RNA Direct Protokoll generierte RNA-Sequenzierbibliotheken wurden auf der Illumina HiSeq 4000 Plattform sequenziert. Die Bibliotheksstatistik wurde nach einem Downsampling auf  $2 \times 10^7$  Reads mithilfe des Picard RNA-Analysetools generiert. Eine MBC-korrigierte Statistik wurde anhand der gleichen Pipeline generiert, nachdem MBCs mit UmiAwareMarkDuplicatesWithMateCigar markiert wurden. (A) Frisch gefrorener („intakter“) RNA-Input; (B) FFPE-Input.

**A. Universelle Human-Referenz-RNA (UHRR)**

Protokoll	SureSelect XT HS RNA				SureSelect XT RNA Direct			
	100	100	10	10	100	100	10	10
Probeneinsatz (ng)	100	100	10	10	100	100	10	10
Analysierte Reads (Mio.)	1,9	1,9	1,9	1,9	1,9	1,9	1,9	1,9
rRNA-Rate (%)	0,1 %	0,1 %	0,1 %	0,1 %	0,1 %	0,1 %	0,1 %	0,1 %
Strangspezifität (%)	99,1 %	99,1 %	99,0 %	99,0 %	99,1 %	99,2 %	99,1 %	99,1 %
Duplikationsrate (%)	17,3 %	16,7 %	24,3 %	24,7 %	17,9 %	20,0 %	35,8 %	38,7 %
Geschätzte Bibliotheksgröße ( $\times 10^6$ )	24,2	25,2	16,2	15,9	27,5	23,6	10,3	9,2
MBC-korrigierte Duplikationsrate (%)	7,1 %	6,4 %	17,1 %	17,8 %	n. z.	n. z.	n. z.	n. z.
MBC-korrigierte geschätzte Bibliotheksgröße ( $\times 10^6$ )	65,9	72,8	24,6	23,6	n. z.	n. z.	n. z.	n. z.

**B. Seraseq FFPE Tumor Fusion RNA Reference Material v2 (SeraCare)**

Protokoll	SureSelect XT HS RNA				SureSelect XT RNA Direct			
	100	100	10	10	100	100	10	10
Probeneinsatz (ng)	100	100	10	10	100	100	10	10
Analysierte Reads (Mio.)	18,6	18,6	18,7	18,6	19	18,8	18,9	18,9
rRNA-Rate (%)	0,1 %	0,2 %	0,1 %	0,1 %	0,1 %	0,2 %	0,1 %	0,1 %
Strangspezifität (%)	98,6 %	98,6 %	98,6 %	98,7 %	98,9 %	98,9 %	98,9 %	98,9 %
Duplikationsrate (%)	24,8 %	24,7 %	36,0 %	45,0 %	23,5 %	29,4 %	69,1 %	67,2 %
Geschätzte Bibliotheksgröße ( $\times 10^6$ )	15,5	15,5	9,6	6,9	18,6	13,3	3	3,3
MBC-korrigierte Duplikationsrate (%)	13,2 %	13,7 %	29,3 %	38,8 %	n. z.	n. z.	n. z.	n. z.
MBC-korrigierte geschätzte Bibliotheksgröße ( $\times 10^6$ )	32,5	31,2	12,7	8,66	n. z.	n. z.	n. z.	n. z.

## Einfluss eindeutiger molekularer Identifikatoren (MBC) auf RNA-Sequenzierbibliotheken

Wie bereits angesprochen, ist die exakte Bestimmung von Bibliothekskomplexitäten für die Bewertung der Effizienz eines Verfahrens zur Bibliothekserstellung von wesentlicher Bedeutung. Die oben angegebenen Bibliothekskomplexitätsschätzungen basieren auf der Annahme, dass es sich bei Read-Paaren mit identischem Start und Stopp um PCR-Duplikate handelt, die von dem gleichen ursprünglichen Bibliotheksmolekül abstammen. Duplikate können allerdings auch durch zufällige Fragmentierung entstehen, die in der Bildung zweier unabhängiger Fragmente mit identischen Enden resultiert. Da Fragmentierungsduplikate tatsächlich unabhängige Bestandteile der Sequenzierbibliothek sind, sollte die Bestimmung von Bibliothekskomplexitäten nur auf der Grundlage echter PCR-Duplikate erfolgen.

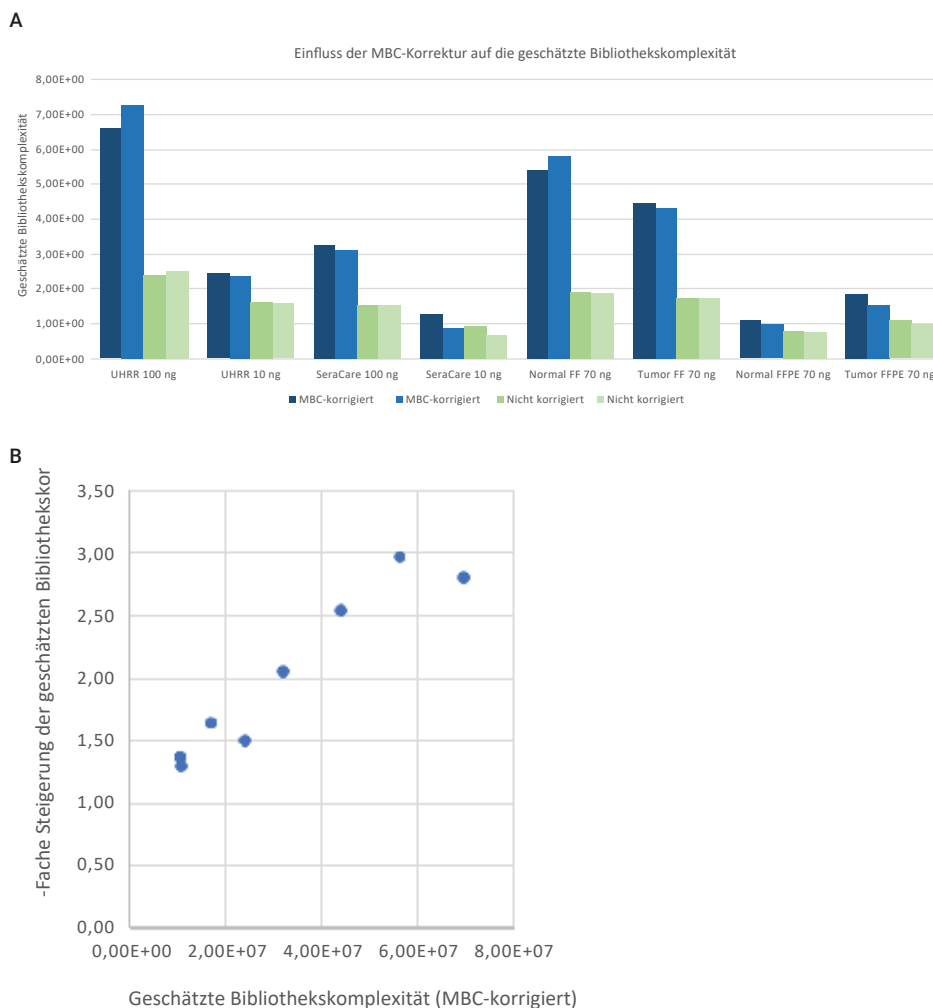
Da durch die starke Expression einiger Gene die Zufallschance der Entstehung von Fragmentierungsduplikaten steigt, entsteht bei der RNA-Sequenzierung möglicherweise eine höhere Zahl von Fragmentierungsduplikaten als bei der DNA-Sequenzierung. Dies wiederum führt zu ungenaueren Schätzungen der Bibliothekskomplexität. Ein möglicher Vorteil des RNA XT HS Arbeitsablaufs besteht darin, dass Bibliotheken mit XT HS Adaptoren erstellt werden, die einzelne 10-bp-MBCs enthalten. Unsere Hypothese war, dass diese MBCs eine Unterscheidung zwischen PCR- und Fragmentierungsduplikaten ermöglichen könnten. Wir wiederholten die Analyse der oben gezeigten RNA XT HS Sequenzierdaten unter Anwendung einer modifizierten Datenanalysepipeline, bei der MBC-Daten genutzt werden (Hinweis: RNA Direct Bibliotheken enthalten keinen MBC und wurden daher aus der Wiederholungsanalyse ausgeschlossen). Die Ergebnisse dieser Analyse sind in Tabelle 1 und Abbildung 4 dargestellt.

Wie Abbildung 4 zeigt, führt die Bereinigung um Fragmentierungsduplikate zu höheren Schätzwerten der entsprechenden Bibliothekskomplexitäten. Insbesondere bei größeren Einsatzmengen wird die Bibliothekskomplexität um das etwa Dreifache unterschätzt, wenn

Fragmentierungsduplikate nicht berücksichtigt werden. Das größere Bias für umfangreichere Bibliotheken und die entsprechend höheren Einsatzmengen sind zu erwarten, da die Wahrscheinlichkeit der Beobachtung von Fragmentierungsduplikaten größer ist, je mehr von der gleichen codierenden Sequenz abgeleitete Moleküle in eine Bibliothek verarbeitet werden.

Unser nächstes Ziel war die Untersuchung der Auswirkungen einer MBC-Korrektur auf „reale Proben“. Dazu erstellten wir RNA XT HS Bibliotheken aus zueinander gehörendem Brusttumor- und angrenzendem Normalgewebe, das jeweils sowohl in frisch

gefrorener Form als auch als FFPE-Probe vorlag. Diese Bibliotheken wurden mit den UHRR und SeraCare RNA XT HS Bibliotheken angereichert, sequenziert und analysiert, wie oben beschrieben. Die zusammenfassende Statistik für diese Bibliotheken ist Tabelle 2 und Abbildung 4 zu entnehmen. Für die Sequenzdaten aus Tumor- und Normalproben, sowohl als frisch gefrorene als auch als FFPE-Proben, sehen wir eine ausgezeichnete Strangspezifität und rRNA-Rate, vergleichbar mit unseren UHRR und SeraCare Bibliotheken. Auffällig ist zudem, dass die FFPE-Bibliotheken signifikant kleiner sind als die aus frisch gefrorenen Proben.



**Abbildung 4.** Einfluss der Korrektur mittels eindeutiger molekularer Identifikatoren (MBC) auf die geschätzten Bibliotheksgrößen. Die Größen der RNA XT HS Bibliotheken wurden entweder durch Picard-Deduplizierung (nicht korrigiert) oder nach der Verwendung von MBCs (MBC-korrigiert) berechnet. Bei der Picard-Deduplizierung wird angenommen, dass es sich bei allen Duplikaten um PCR-Duplikate handelt, während die MBC-Korrektur zwischen PCR- und Fragmentierungsduplikaten unterscheidet. **A)** Geschätzte Bibliothekskomplexitäten für mehrere aus verschiedenen Proben und Einsatzmengen generierte RNA XT HS Sequenzierbibliotheken. **B)** Die Unterschiede zwischen den MBC-korrigierten und den nicht korrigierten Bibliothekskomplexitäten fallen für größere Bibliotheken und dementsprechend höhere Einsatzmengen deutlicher aus.

Dieses Phänomen ist gut dokumentiert. Bei Analyse unserer FFPE-Sequenzdaten unter Anwendung des MBC-korrigierten Verfahrens ergaben sich Verbesserungen der geschätzten Bibliothekskomplexität und damit der gewonnenen Daten um das 0,3- – 3,5-Fache. Diese Daten wären verloren gegangen, hätten wir angenommen, dass es sich bei allen Duplikaten ausschließlich um PCR-Duplikate und nicht zum Teil auch um Fragmentierungsduplikate handelt. Dies demonstriert, dass die genaue Identifizierung von PCR- und Fragmentierungsduplikaten in diesen Proben ein kritischer Schritt ist, um die aus FFPE-RNA-Proben gewonnenen Sequenzierinformationen zu maximieren.

### Detektion von Genfusionen

Die gezielte RNA-Seq hat sich als nützlich für die Detektion von Genfusionen erwiesen, insbesondere bei schwierigen Proben (FFPE). Wir wollten ermitteln, ob unser RNA XT HS Ansatz unter den Bedingungen einer Anreicherung mit Exomsonden (im Vergleich zu einem fusionsspezifischen Panel) und einer Maximierung der Sequenzlänge für diesen Anwendungsbereich geeignet ist. Wie zuvor erwähnt, wurden unsere RNA XT HS Bibliotheken aus allen Proben (UHRR, SeraCare, frisch gefroren und FFPE) einer gezielten Anreicherung unter Anwendung des SureSelect Human All Exon V7 Exoms unterzogen, und die angereicherten Bibliotheken wurden bei einer Read-Länge von  $2 \times 150$  mittels Paired-End-Sequenzierung sequenziert. Die Sequenzdaten aus frisch gefrorenen bzw. FFPE-Tumor-/Normalproben wurden mit STAR-Fusion analysiert und mit FusionInspector visualisiert.

In FusionInspector wurde eine Reihe mutmaßlicher Fusionen detektiert, darunter bekannte Falschpositive wie VDJ-Rekombinationsereignisse, welche auf das Vorliegen von Immunzellen in einigen Proben hindeuten. Die VDJ-Rekombinationsereignisse wurden aus Tabelle 3 entfernt. Aus der Datenanalyse ergeben sich 15 potenzielle Genfusionen in unseren Proben, wobei für 12 der 15 Fusionspartner bereits eine Beteiligung an tumorbedingten Genfusionen berichtet worden war, wenn auch nicht in der beobachteten Kombination. Eine der Fusionen, FCHSD2-FAM168A (in Tabelle 3 grau hervorgehoben), ist in der Vergangenheit mit Brustkrebs in Verbindung gebracht worden.

**Tabelle 2.** Globale Sequenzierstatistik für Agilent SureSelect RNA XT HS Bibliotheken aus frisch gefrorenen und FFPE-Proben. RNA-Sequenzierbibliotheken wurden gemäß dem RNA XT HS Protokoll aus zueinander gehörendem Tumor- und angrenzendem Normalgewebe in Form von frisch gefrorenen und FFPE-Proben generiert. Die Bibliotheken wurden auf der Illumina HiSeq 4000 Plattform sequenziert. Nach einem Downsampling auf  $2 \times 10^7$  Reads wurde mithilfe des Picard RNA-Analysertools eine Bibliotheksstatistik erstellt, entweder mit oder ohne MBC-Markierung.

Probentyp	Frisch gefroren				FFPE			
	Normal	Normal	Tumor	Tumor	Normal	Normal	Tumor	Tumor
Probeneinsatz (ng)	70	70	70	70	70	70	70	70
Analysierte Reads (Mio.)	18,8	18,8	18,7	18,8	18,6	18,6	18,7	18,7
rRNA-Rate (%)	0,2 %	0,2 %	0,2 %	0,2 %	0,2 %	0,2 %	0,1 %	0,1 %
Strangspezifität (%)	99,0 %	99,0 %	99,1 %	99,2 %	98,4 %	98,4 %	98,8 %	98,8 %
Duplikationsrate (%)	21,0 %	21,0 %	23,0 %	23,0 %	41,0 %	43,0 %	33,0 %	35,0 %
Geschätzte Bibliotheksgröße ( $\times 10^6$ )	19	18,8	17,3	17,3	7,9	7,4	10,9	9,8
MBC-korrigierte Duplikationsrate (%)	8,0 %	8,0 %	10,0 %	10,0 %	33,0 %	35,0 %	22,0 %	25,0 %
MBC-korrigierte geschätzte Bibliotheksgröße ( $\times 10^6$ )	53,9	58,2	44,7	43,1	11	9,93	18,6	15,4

**Tabelle 3.** Detektion von Genfusionen in frisch gefrorenen und FFPE-Normal-/Tumorproben. RNA XT HS Sequenzierbibliotheken wurden sequenziert und anschließend unter Anwendung einer STAR-Fusion-Pipeline auf Genfusionen analysiert. Die mutmaßlichen Fusionen, die in einer oder mehreren Proben detektiert wurden, sind in den einzelnen Zeilen jeweils zusammen mit unterstützenden Hinweisen aufgeführt. Junction- und Support-Reads sind als Reads pro Million eindeutiger Bibliotheks-Read-Paare angegeben. Jede Spalte repräsentiert den Durchschnitt zweier technischer Replikate. Die in Grau hervorgehobene Fusion ist in der TGCA-Datenbank als eine mit Brustkrebs assoziierte Fusion angegeben; bei ihr könnte es sich um eine Treibermutation für diesen Tumor handeln. Beobachtete VDJ-Rekombinationsereignisse, die auf Immunzellen in den Proben hindeuten, wurden von der Analyse ausgeschlossen. Die Detektion basiert auf der Analyse von  $0,7 - 1,4 \times 10^7$  eindeutiger Read-Paare pro Probe.

Mutmaßliche Fusion	Frisch gefroren				FFPE			
	Normal		Tumor		Normal		Tumor	
	Junction	Support	Junction	Support	Junction	Support	Junction	Support
MYLK-LPAR6	0,00	0,00	8,86	0,86	0,00	0,00	1,68	0,00
FCHSD2-FAM168A	0,00	0,04	5,60	0,21	0,00	0,00	2,79	0,04
FAM157A-RB1	0,00	0,00	2,59	0,04	0,00	0,00	0,91	0,00
RP4-565E6.1-HYDIN	0,00	0,00	1,47	0,00	0,00	0,00	1,33	0,00
CDR2-FRG1	0,00	0,00	0,42	0,00	0,00	0,00	0,48	0,00
SLC7A5-RP11-645C24.2	0,00	0,04	0,33	0,20	0,00	0,00	0,09	0,04
PHRF1-TXNDC5	0,00	0,00	2,78	0,17	0,00	0,00	0,00	0,00
NUDT1-AC004840.8	0,00	0,00	1,71	0,00	0,00	0,00	0,00	0,00
MPZL1-RCS1	0,00	0,00	1,34	0,06	0,00	0,00	0,00	0,00
CCDC66-SLMAP	0,00	0,00	0,96	0,00	0,00	0,00	0,00	0,00
UBE2Q2-C15orf27	0,00	0,00	0,77	0,06	0,00	0,00	0,00	0,00
POLR2J-AC004980.9	0,00	0,00	0,00	0,00	11,53	0,00	0,00	0,00
POLR2J-UPK3B	0,00	0,00	0,00	0,00	3,96	0,00	0,00	0,00
RP11-634B7.4-TRIM58	0,36	0,00	0,00	0,00	0,00	0,00	0,00	0,00
RB1-MYLK-AS1	0,00	0,00	0,16	0,03	0,00	0,00	0,00	0,00

Möglicherweise handelt es sich um eine Treibermutation für diesen Tumor<sup>5</sup>. Für alle anderen tumorspezifischen Fusionen ist mindestens ein Fusionspartner (meist jedoch beide) in der TCGA-Datenbank aufgeführt und mit der Tumorentstehung assoziiert. Für ein Fusionstranskript A:B stellen wir fest, dass sowohl A als auch B in der Tumordatenbank gelistet sind. Es gibt allerdings keinen Eintrag für A:B. Stattdessen sind Beispiele für A:C- und D:B-Fusionstranskripte vorhanden. Diese Daten deuten darauf hin, dass mit einem Exom angereicherte RNA XT HS Bibliotheken es Forschern selbst bei schwierigen Proben ermöglichen, Fusionen auf der Ebene des Gesamt-Transkriptoms zu identifizieren.

## Schlussfolgerung

Mit dem Ziel, eine effizientere Konstruktion gezielter RNA-Seq-Bibliotheken zu ermöglichen, arbeiteten wir unter Verwendung von Komponenten des SureSelect XT RNA Direct Kits zur Bibliothekserstellung zusammen mit dem SureSelect XT HS Kit zur Target-Anreicherung an einer Verbesserung des aktuell gebräuchlichen SureSelect XT RNA Direct Arbeitsablaufs. Der neue SureSelect XT HS RNA Arbeitsablauf weist mehrere signifikante Verbesserungen auf:

- Die Lyophilisierung des RNA-Inputs wurde durch die direkte Zugabe von Fragmentierungspuffer zur RNA-Probe ersetzt.
- Der Aufreinigungsschritt mit SPRI-Beads nach der Synthese des ersten cDNA-Strangs wurde eliminiert.
- Ein alternatives PCR-Enzym, das uracilhaltige Templates nicht amplifiziert, wurde verwendet, um die UDG-Behandlung zur Aufrechterhaltung der Strangspezifität überflüssig zu machen.
- Das Capture-Verfahren mittels Sonde („Bait“) wurde von einer 24-Stunden-Hybridisierung auf ein 90-minütiges Schnellhybridisierungsprotokoll umgestellt.

- Die XT Sequenzieradapter wurden durch XT HS Adapter ersetzt, welche die Nutzung von MBCs für eine präzisere Schätzung der Bibliothekskomplexität und „Wiederfindung“ der Reads ermöglichen.

Wir stellten fest, dass sich mit dem RNA XT HS Arbeitsablauf die Bearbeitungszeit für Bibliothekserstellung und Target-Anreicherung von 2–3 Tagen auf 1–2 Tage reduziert. Die mit dem neuen, optimierten Arbeitsablauf generierten RNA-Seq-Bibliotheken sind bei hohen Einsatzmengen im Hinblick auf Strangspezifität, rRNA-Raten und Mapping-Raten nicht von Bibliotheken zu unterscheiden, die nach dem SureSelect XT RNA Direct Protokoll erstellt wurden. Bei geringeren Einsatzmengen (10 ng) beeinträchtigte der verkürzte Arbeitsablauf die Gesamtleistung nicht und erwies sich sogar als effizienter. Wir fanden auch heraus, dass die Einbeziehung von MBCs in das verbesserte Adapterdesign (XT HS) nun die Identifizierung von Fragmentierungsduplikaten ermöglicht. Durch diese verbesserte Identifizierung von Fragmentierungsduplikaten verbessern sich die Sequenzierergebnisse insgesamt, da ein Verlust von Reads vermieden wird, der bei Anwendung standardmäßiger Start-Stopp-Duplikationsmethoden auftreten würde.

Bei der Untersuchung „realer Proben“ lieferte der RNA XT HS Arbeitsablauf weiterhin Daten von hoher Qualität, selbst für FFPE-Proben. In einer vorläufigen Genfusionsanalyse wurden potenzielle Genfusionen in unseren frisch gefrorenen und FFPE-Tumorproben detektiert – ein Hinweis auf einen möglichen praktischen Anwendungsbereich der Sequenzierung von FFPE-RNA mittels RNA XT HS. Obgleich wir hier keine Genexpressions- oder Spleißdaten präsentiert haben, deutet unsere vorläufige Analyse darauf hin, dass die gezielte RNA-Seq für die globale Genexpressionsanalyse, die Detektion von Spleißvarianten und abweichenden Expressionen sowie die allelspezifische Expressionsanalyse verwendet werden kann. Schließlich ist

der SureSelect RNA XT HS Arbeitsablauf auf das DNA SureSelect XT HS Kit zur Target-Anreicherung „abgestimmt“, was die parallele Sequenzierung von DNA und RNA ausgehend von der gleichen Probe erleichtert. Dies könnte für zahlreiche Applikationen einen Vorteil darstellen, etwa in der Multiomik-Forschung oder bei der parallelen Verarbeitung von Proben für die TMB-MSI-Analyse und Fusionsdetektion.

## Abkürzungen

FFPE, formalinfixiert, Paraffin-eingebettet; TPM, Transkripte pro Million Kilobasen; nt, Nukleotid; UHRR, universelle Human-Referenz-RNA; MBC, molekularer Barcode; UDG, Uracil-Deglycosylase

## Literatur

- 1) Cieslik, M., *et al.* The Use of Exome Capture RNA-Seq for Highly Degraded RNA with Application to Clinical Cancer Sequencing. *Genome Res.* **2015**, 25, 1372–1381.
- 2) Mittempergher, L., *et al.*, MammaPrint and Blueprint Molecular Diagnostics Using Targeted RNA Next-Generation Sequencing Technology. *J. Mol. Diagn.* **2019**, 21, 808–823.
- 3) Jones, J. C.; Alex Siebold, A.; Lucas, A. B. SureSelect XT RNA Direct Protocol Provides Simultaneous Transcriptome Enrichment and Ribosomal Depletion of FFPE RNA. Agilent Technologies application note, publication number 5991-8119EN, **2017**.
- 4) Haas, B., *et al.* STAR-Fusion: Fast and Accurate Fusion Transcript Detection from RNA-Seq. *bioRxiv.* 120295 (**2017**).
- 5) Hu, X., *et al.* TumorFusions: An Integrative Resource for Cancer-Associated Transcript Fusions. *Nucleic Acids Res.* **2018**, 4, 46(D1), D1144-D1149.

## Anhang

### 1. Detailliertes Protokoll für die Konstruktion von RNA-Sequenzierbibliotheken

#### RNA-Probenvorbereitung

Gesamt-RNA aus FFPE-Schnitten wurde mit einem Qiagen RNeasy FFPE Kit gemäß den Anweisungen des Herstellers isoliert. RNA aus gefrorenem Gewebe wurde mit einem Qiagen RNeasy Mini Kit isoliert. Alle Gesamt-RNA-Proben wurden auf einem Agilent 2100 Bioanalyzer System mit dem RNA 6000 Pico Kit analysiert. Vor dem Laden auf den Chip wurden die Proben für 2 Minuten auf 80 °C erhitzt. RIN- und DV200-Werte wurden mit der Bioanalyzer Software berechnet. Diese Qualitätsmetriken für alle getesteten Gesamt-RNA-Proben sind in Tabelle S1 aufgeführt.

#### Generierung von cDNA mit Reagenzien aus dem Agilent SureSelect XT RNA Direct Kit

*Hinweis: Eine Stammlösung von 4 µg/µl Actinomycin D in DMSO wurde im Vorfeld hergestellt und in Aliquoten für die Einmalverwendung (3 µl) bei -20 °C aufbewahrt.*

*Hinweis: Fragmentierungsmischung, Mastermix für den ersten Strang und Enzym- und Oligomischungen für den zweiten Strang aus dem SureSelect XT RNA Direct Kit wurden auf Eis aufgetaut, für 5 Sekunden bei hoher Geschwindigkeit gevortext und kurz herunterzentrifugiert.*

1. Gesamt-RNA-Proben wurden in 4 µl nukleasefreiem Wasser angesetzt. Die Einsatzmengen unterschieden sich je nach Experiment und sind im Abschnitt „Ergebnisse“ angegeben.

*Hinweis: Geringere Volumen für den RNA-Input (< 4 µl) sind möglich, aber größere RNA-Probenvolumen (> 4 µl) werden nicht empfohlen.*

2. Die RNA-Probe wurde mit Fragmentierungsmischung auf ein Endvolumen von 20 µl aufgefüllt.
3. Die RNA-Probe wurde durch Erhitzen in einem SureCycler 8800 (oder vergleichbarer Thermocycler) fragmentiert. Die Bedingungen wurden anhand der Qualitätsmetriken

**Tabelle S1.** Qualitätsmetriken der Proben und Fragmentierungsbedingungen.

Probenbeschreibung	RIN	DV200	Fragmentierung
Universelle Human-Referenz-RNA (UHRR)	9,2	94 %	94 °C, 8 Min.
Seraseq FFPE Tumor Fusion RNA Reference Material v2	2	54 %	94 °C, 3 Min., 65 °C, 2 Min.
Brust, normal, gefroren	6,3	94 %	94 °C, 8 Min.
Brust, Tumor, gefroren	5	88 %	94 °C, 8 Min.
Brust, normal, FFPE	2,1	48 %	65 °C, 5 Min.
Brust, Tumor, FFPE	2	47 %	65 °C, 5 Min.

**Tabelle S2.** Reaktionsmischung für den ersten Strang.

Reagenz	Volumen für eine Reaktion	Volumen für acht Reaktionen + 10 %Überschuss
Mastermix für RNA-Sequenzierung des ersten Strangs	8 µl	70,4 µl
Actinomycin-D (120 ng/µl)	0,5 µl	4,4 µl
<b>Gesamt</b>	<b>8,5 µl</b>	<b>74,8 µl</b>

**Tabelle S3.** Parameter für das SPRI-Waschprotokoll.

Volumen AMPure Beads	105 µl (1,8x Volumen)
Bead-Inkubationszeit	5 Minuten
Waschschritt mit 70 % Ethanol (zweimal)	200 µl
Trockenzeit bei 37 °C	1–2 Min. oder weniger
Elutionsvolumen	50 µl nukleasefreies Wasser

4. Eine Stammlösung von 4 µg/µl Actinomycin-D in DMSO wurde in Wasser auf 120 ng/µl verdünnt (3 µl Actinomycin-D + 97 µl Wasser).
5. Für die Synthese des ersten Strangs wurde eine Master-Reaktionsmischung für acht Proben erstellt (Tabelle S2). Die Reaktionsmischung wurde gevortext und bis zum Gebrauch auf Eis gelagert.
6. Jeder fragmentierten 20-µl-Probe auf Eis wurden 8,5 µl der Reaktionsmischung für den ersten Strang hinzugefügt. Die Proben wurden durch Vortexen gemischt und kurz herunterzentrifugiert.
7. Die 28,5-µl-Reaktionen wurden in einem vorprogrammierten SureCycler 8800 für 10 Minuten bei 25 °C und dann für 40 Minuten bei 37 °C inkubiert und anschließend bis zur Synthese des zweiten Strangs bei 4 °C (oder auf Eis) gelagert.
8. Die Röhrchen mit der Enzymmischung für den zweiten Strang und die Endreparatur sowie der Oligomischung für den zweiten Strang und die Endreparatur wurden jeweils vor Verwendung gevortext.
9. Jeweils 25 µl der Enzymmischung für den zweiten Strang und die Endreparatur (blauer Deckel) wurden der 28,5-µl-Reaktion für den ersten Strang auf Eis hinzugegeben.
10. Unmittelbar danach wurden je 5 µl der Oligomischung für den zweiten Strang und die Endreparatur (gelber Deckel) hinzugefügt.

11. Die Proben wurden verschlossen, gevortext, kurz herunterzentrifugiert und dann auf Eis zurückgestellt. 12. Die 58,5- $\mu$ l-Reaktionen wurden in einem vorprogrammierten SureCycler 8800 für 60 Minuten bei 16 °C inkubiert und anschließend bis zum SPRI-Aufreinigungsschritt bei 4 °C (oder auf Eis) gelagert.
13. Nach der Synthese des zweiten Strangs wurde die cDNA gemäß dem in Tabelle S3 dargestellten Protokoll mittels SPRI aufgereinigt.

**Erstellung von SureSelect XT HS cDNA-Bibliotheken unter Anwendung von Reagenzien aus dem SureSelect XT HS Kit zur Bibliothekserstellung**

14. Endreparatur, dA-Tailing und XT HS Adapterligation der cDNA wurden unter Anwendung der Reagenzien und Anweisungen aus Kit und Protokoll zum SureSelect XT HS System zur Target-Anreicherung für die Illumina Paired-End-Multiplex-Sequenzierung (G9702) durchgeführt.

*Hinweis: Das SureSelect XT HS Protokoll wurde ab Schritt 3 auf Seite 27 befolgt, mit den in den Schritten 15 bis 20 dieses Protokolls beschriebenen Ausnahmen und Modifikationen.*

15. Zu Schritt 6 auf Seite 34 des SureSelect XT HS Protokolls. Die Precapture-PCR-Amplifikation der Bibliotheken wurde gemäß den Anweisungen des Protokolls durchgeführt. Die Anzahl der PCR-Zyklen wurde folgendermaßen modifiziert:
  - a. UHRR, Seraseq™ v2 für 12 PCR-Zyklen (RNA von hoher Qualität)
  - b. Proben aus Brustgewebeblöcken für 14 PCR-Zyklen (RNA von geringer Qualität)
16. Ausbeute und Verteilung der Molekülgrößen für die PreCap-Bibliotheken wurden mittels Bioanalyzer oder TapeStation untersucht.

**Target-Anreicherung und Sequenzierung der SureSelect XT HS cDNA-Bibliotheken**

17. Zu Schritt 1 auf Seite 46 des SureSelect XT HS Protokolls. Die Target-Anreicherung wurde mit 200 ng PreCap-Bibliothek als Input durchgeführt. Dazu wurden SureSelect XT HS Reagenzien für eine Schnellhybridisierung verwendet. Für die Hybridisierung wurden je 5  $\mu$ l SureSelect Human All Exon V7 Exomsonden verwendet.

18. Die Bindung mit Streptavidin-Beads und die anschließenden Waschschrte wurden gemäß Protokoll durchgeführt.
19. Die Postcapture-PCR-Amplifikation wurde gemäß Protokoll mit je 12 PCR-Zyklen pro Probe durchgeführt. Ausbeute und Verteilung der Molekülgrößen für die PostCap-Bibliotheken wurden mittels Bioanalyzer oder TapeStation untersucht.
20. Alle Bibliotheken wurden auf einer Illumina HiSeq 4000 Plattform bei einer Read-Länge von 2  $\times$  150 sequenziert.

[www.agilent.com](http://www.agilent.com)

**Ausschließlich zu Forschungszwecken.  
Nicht für Diagnoseverfahren geeignet.**

Änderungen vorbehalten.

PR7000-2381  
© Agilent Technologies, Inc. 2019, 2020  
Gedruckt in den USA, Mittwoch, 29. Januar 2020  
5994-1644DEE

